(19) Europäisches Patentamt
European Patent Offi
Offi européen des br vets

(11) EP 1 081 585 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
07.03.2001 Bulletin 2001/10

(51) Int Cl.[7]: **G06F 3/06**

(21) Application number: **00307288.1**

(22) Date of filing: **24.08.2000**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **31.08.1999 US 386252**

(71) Applicant: **International Business Machines Corporation**
Armonk, NY 10504 (US)

(72) Inventors:
- **Goodgold, Stuart R., IBM UK Ltd., I.P.L.**
  Winchester, Hampshire SO21 2JN (GB)
- **Azevedo, Ruth Enid, IBM UK Ltd., I.P.L.**
  Winchester, Hampshire SO21 2JN (GB)
- **McNutt, Bruce, IBM UK Ltd., I.P.L.**
  Winchester, Hampshire SO21 2JN (GB)

(74) Representative: **Burt, Roger James, Dr.**
IBM United Kingdom Limited
Intellectual Property Department
Hursley Park
Winchester Hampshire SO21 2JN (GB)

(54) Method and system for reporting disk utilisations for disk subsystems

(57) The present invention provides a method and system for reporting disk utilisation in a computer system (200). The method includes collecting at least one parameter pertaining to a disk array (206) in a disk subsystem (202) by a processor (210) in the disk subsystem (202); transferring the at least one parameter from the processor (210) to an open system host (204); and calculating disk utilisation based upon the at least one parameter. The method and system is able to provide disk utilisation which is understandable to the open system host (204). In a preferred embodiment, the disk utilisation is provided by collecting the relevant parameters from the microcode in the processor (210) of the disk subsystem (202), and then calculating the average disk utilisation per disk based upon the parameters. Thus, an open system host (204) is able to obtain a useful characteristic for determining the performance of the disk subsystem.

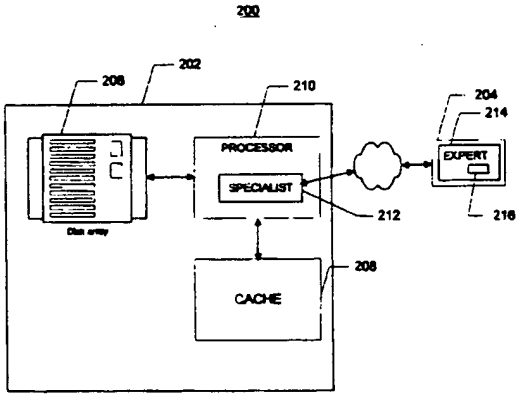


FIG. 2

EP 1 081 585 A1

Description

[0001] The present invention relates to computer systems, and more specifically to disk subsystems in computer systems.

[0002] Mainframe disk subsystems are well known in the art. Figure 1 illustrates a mainframe computer system with a conventional disk subsystem. The computer system 100 includes a disk subsystem 102 and a host 104. The disk subsystem 102 includes a disk array 106, a cache 108, and a processor 110 which controls the functioning of the disk subsystem 102. The host 104 comprises software specifically for communication with the disk subsystem 102, and does so by interfacing with the processor 110. One of the functions of the processor 110 is to provide statistics and other characteristics of the disk array 106 to the host 104. Examples of disk characteristics of interest to the host 104 include a table of ownership, the amount of space used on the disks 106, and performance characteristics. An important performance characteristic is disk utilisation, which tells the host 104 how busy each disk is in the array 106. In a mainframe environment, as illustrated in Figure 1, obtaining the disk utilisation characteristic is well known in the art. However, a problem arises when the disk subsystem attempts to communicate with open system hosts, such as hosts which run UNIX, Windows NT, or some other type of open operating system. The parameters relevant to disk characteristics sent from the processor 110 to the host 104 is not in a format which is understandable by the open system host. Thus, in an open systems environment, the open system host is unable to obtain disk characteristics, including the disk utilisation characteristic.

[0003] Accordingly, there exists a need for a method and system for providing parameters for a disk subsystem to an open system host. The method and system should be able to provide disk utilisation which is understandable to the open system host. The present invention seeks to address such a need.

[0004] According to one aspect of the present invention, a method is provided, for reporting disk utilisation in a computer system, comprising the steps of:

(a) collecting at least one parameter pertaining to a disk array in a disk subsystem by a processor in the disk subsystem;
(b) transferring the at least one parameter from the processor to an open system host; and (c) calculating disk utilisation based upon the at least one parameter.

[0005] According to a second aspect of the present invention, a computer system is provided, comprising: a disk subsystem, comprising: a disk array, and a processor coupled to the disk array, the processor being capable of collecting at least one parameter pertaining to the disk array and transferring the at least one parameter to an open system host; and the open system host coupled to the disk subsystem, the open system host being capable of receiving the at least one parameter and calculating a disk utilisation based upon the at least one parameter.

[0006] According to a third aspect of the present invention, a computer readable medium is provided with computer instructions for reporting disk utilisation in a computer system, the instructions being for:
(a) collecting at least one parameter pertaining to a disk array in a disk subsystem by a processor in the disk subsystem; (b) transferring the at least one parameter from the processor to an open system host; and (c) calculating disk utilisation based upon the at least one parameter.

[0007] The present invention provides a method and system for reporting disk utilisation in a computer system. The method includes collecting at least one parameter pertaining to a disk array in a disk subsystem by a processor in the disk subsystem; transferring the at least one parameter from the processor to an open system host; and calculating disk utilisation based upon the at least one parameter. The method and system is able to provide disk utilisation which is understandable to the open system host.

[0008] In a preferred embodiment, the disk utilisation is provided by collecting the relevant parameters from the microcode in the processor of the disk subsystem, and then calculating the average disk utilisation per disk based upon the parameters. Thus, an open system host is able to obtain a useful characteristic for determining the performance of the disk subsystem. Preferably, the open system host comprises a database for storing the at least one parameter.

[0009] One embodiment of a method for reporting disk utilisation in a computer system according to the present invention, comprises the steps of:

(a) collecting at least one parameter pertaining to a disk array in a disk subsystem by a specialist in a processor in the disk subsystem;
(b) transferring the at least one parameter from the specialist to an expert in an open system host; and
(c) calculating disk utilisation by the expert based upon the at least one parameter.

[0010] One embodiment of a computer system according to the invention, comprises: a disk subsystem, comprising: a disk array; and a processor coupled to the disk array, comprising a specialist for collecting at least one parameter pertaining to the disk array and transferring the at least one parameter to an open system host; and the open system host coupled to the disk subsystem, comprising an expert for receiving the at least one parameter and calculating a disk utilisation based upon the at least one parameter.

[0011] Yet a further embodiment of a method according to the invention for reporting disk utilisation in a com-

put r system, comprises th st ps of: (a) requesting at least one disk characteristic pertaining to a disk array in a disk subsystem by a user at an open system host; (b) s nding a request for th at least one param ter from an expert in th open syst m host to a specialist in a processor in the disk subsystem, wherein the at least one parameter is relevant to the at least one disk characteristic;

(c) collecting the at least one parameter by the specialist from microcode in the processor according to the request; (d) transferring the at least one parameter from the specialist to the expert; (e) storing the at least one parameter in a database on the open system host; and (f) calculating a disk utilisation using a disk utilisation routine in the expert.

[0012] The disk subsystem may be an Enterprise Storage Server (ESS) disk subsystem. Also, the open system host may be an UNIX or a Windows system host.

[0013] For a better understanding of the present invention reference will now be made to the accompanying drawings, in which, by way of example:

Figure 1 illustrates a mainframe computer system with a conventional disk subsystem;

Figure 2 illustrates a preferred embodiment of a computer system with a disk subsystem in accordance with the present invention;

Figure 3 is a flow chart illustrating a preferred embodiment of reporting disk utilisation in accordance with the present invention; and

Figure 4 is a flow chart illustrating in more detail the reporting of disk utilisation in accordance with the present invention.

[0014] The present invention provides a method and system for providing parameters for a disk subsystem to an open system host. The following description is presented to enable one of ordinary skill in the art to make and use the invention and is provided in the context of a patent application and its requirements. Various modifications to the preferred embodiment will be readily apparent to those skilled in the art and the generic principles herein may be applied to other embodiments. Thus, the present invention is not intended to be limited to the embodiment shown but is to be accorded the widest scope of the appended claims consistent with the principles and features described herein.

[0015] Figure 2 of the drawings shows a computer system 200, which comprises a disk subsystem 202 and an open system host 204. Traditional mainframe hosts (not shown) may also communicate with the disk subsystem 202. The disk subsystem 202 comprises a disk array 206, a cache 208, and a processor 210 which controls th functioning f the disk subsystem 202. Th open system host 204 communicates with th disk subsystem 202 by interfacing with th processor 210 via an Intern t connection. Softwar resides in th processor 210 and the open system host 204 to facilitate this communication; the processor 210 comprises the specialist software 212, and the op n system host 204 comprises the expert software 214. Clearly, software functions may be reproduced by hardware configurations and throughout the specification the terms "specialist" and "expert" used alone are intended to cover either software or hardware capable of performing the required functions.

[0016] One of the functions of the specialist 212 is to gather certain parameters from the microcode (not shown) within the processor 210 which pertain to disk characteristics and transfers them to the expert 214 in the open system host 204. One of the functions of the expert 214 is to receive the parameters from the specialist 212 and to generate reports concerning disk characteristics based upon the parameters. The reports may be in tabular or graphical form. For the computer system 200, microcode within the processor 210 does not provide disk utilisation directly. Thus, the expert 214 comprises a software routine 216 which calculates disk utilisation from the parameters provided by the microcode via the specialist 212. An example of a disk subsystem which may be used in the computer system 200 is the ENTERPRISE STORAGE SERVER (ESS) disk subsystem developed by INTERNATIONAL BUSINESS MACHINES CORPORATION. In the ESS disk subsystem, the disk array 206 is a RAID array, which is well known in the art.

[0017] Figure 3 is a flow chart illustrating a preferred embodiment of reporting disk utilisation in accordance with the present invention. First, the specialist 212 in the processor 210 of the disk subsystem 202 collects the parameters, via step 302. Next, the parameters are transferred from the specialist 212 to the expert 214 in the open system host 204, via step 304. Then, the expert 214 calculates the disk utilisation from the parameters, via step 306. In the preferred embodiment, the calculation is performed with a disk utilisation routine 216 in the expert 214. The disk utilisation, along with other disk characteristics, may be displayed by the expert 214 in the form of a report.

[0018] Figure 4 is a flow chart illustrating in more detail the reporting of disk utilisation in accordance with the present invention. First, once the user logs into the open system host 204, the user requests disk characteristics for the disk subsystem 202, via step 402. This request may be in the form of a schedule. For example, the user may request that certain disk characteristics be collected from the disk subsystem 202 every 10 minutes, or every hour. One of the disk characteristics which may be requested by the user is disk utilisation. Next, the expert 214 sends a request to the specialist 212 for the disk charact ristics, via step 404. When the user requests disk utilisation, the xpert 214 request that the param ters in the microcod which are relevant to disk utilisation b collected. For a computer system 200

which uses the ESS disk subsystem, the relevant microcode parameters comprise the following:

T = time interval in seconds
N = number of disks in a RAID array
MR = milliseconds of read time in the time interval
MW = milliseconds of write time in the time interval
W = total number of writes in the time interval
S = number of stride write destages in the time interval

[0019] The specialist 212 then collects the parameters from the microcode according to the request, via step 406. Next, the specialist 212 transfers the parameters to the expert 214, via step 408. Thus if the user requests that disk characteristics be collected every 10 minutes, then at 10 minute intervals, the parameters would be collected from the microcode and sent to the expert 214. Each time the expert 214 receives the parameters, they are stored in a database (not shown) on the open system host 204, via step 410. When the user is ready to view the disk characteristics, the user requests a report from the expert 214, via step 412. The expert 214 then calculates disk utilisation for each disk in the array 206 from the parameters stored in the database, via step 414, using the disk utilisation routine 216. For a computer system 200 which uses an ESS disk subsystem, the formula for the average disk utilisation per disk is as follows:

$$U = 1/(1000*T*N) * [MR + (MW/2W-S) * (4*(W-S) + N*S)]$$

[0020] The formula is derived based upon the fields which are provided by the microcode and their meaning, and an understanding of the internal working of the hardware of the disk subsystem 202. For example, the formula takes into account how a read, write, and stride write destages are performed by the disk subsystem 202, and how the cache 208 and the disk array 206 are involved. For this formula, U is a number between 0 and 1. Once calculated, the expert 214 then displays the report, which includes disk utilisation, via step 416. The report may be in a tabular or graphical form. Disk utilisation for each disk in the array 206 may be displayed as a percentage, representing the amount of time a disk is in use during the interval of time, T.

[0021] The parameters of the microcode described above is part of a proposed standard currently before the Storage Network Industry Association (SNIA). If the parameters are adopted as the standard, then the method and system of the present invention may be used to conform disk utilisation reporting to this standard.

[0022] A method and system for providing parameters for a disk subsystem to an open system host has been disclosed. The method and system is able to provide disk utilisation which is understandable to the open system host. The disk utilisation is provided by collecting the relevant parameters from the microcode in the processor of the disk subsystem, and then calculating the average disk utilisation per disk based upon the parameters. Thus, an open system host is able to obtain a useful characteristic for determining the performance of the disk subsystem.

[0023] Although the present invention has been described in accordance with the embodiments shown, one of ordinary skill in the art will readily recognise that there could be variations to the embodiments and those variations would be within the scope of the appended claims.

Claims

1. A method for reporting disk utilisation in a computer system, comprising the steps of:

   (a) collecting at least one parameter pertaining to a disk array in a disk subsystem by a processor in the disk subsystem;

   (b) transferring the at least one parameter from the processor to an open system host; and

   (c) calculating disk utilisation based upon the at least one parameter.

2. A method according to claim 1, wherein the collecting step (a) comprises:
   (a1) collecting the at least one parameter by a specialist in the processor;

3. A method according to claim 2, wherein the collecting step (a1) comprises:

   (a1i) requesting at least one disk characteristic pertaining to the disk array from a user at the open system host;

   (a1ii) sending a request from an expert in the open system host to the specialist for the at least one parameter, wherein the at least one parameter is relevant to the at least one disk characteristic; and

   (a1iii) collecting the at least one parameter by the specialist from microcode in the processor according to the request.

4. A method according to any preceding claim, wherein the transferring step (b) comprises:
   (b1) transferring the at least one parameter from a specialist in the processor to an expert in the

op n system host.

5. A method acc rding t any preceding claim, wh re-in the calculating step (c) comprises:

(c1) calculating the disk utilisation by an expert in th op n system host based upon the at least one parameter.

6. A method according to claim 5, wherein the calculating step (c1) comprises:

(c1i) receiving the at least one parameter from a specialist in the processor;

(c1ii) storing the at least one parameter in a database on the open system host; and

(c1iii) calculating the disk utilisation using a disk utilisation routine in the expert.

7. A computer system, comprising:

a disk subsystem, comprising:

a disk array, and

a processor coupled to the disk array, the processor capable of collecting at least one parameter pertaining to the disk array and transferring the at least one parameter to an open system host; and

the open system host coupled to the disk subsystem, the open system host capable of receiving the at least one parameter and calculating a disk utilisation based upon the at least one parameter.

8. A system according to claim 7, wherein the processor comprises:

microcode, the microcode comprising the at least one parameter; and

a specialist for collecting the at least one parameter from the microcode and transferring the at least one parameter to the open system host.

9. A system according to claim 7 or 8, wherein the open system host comprises:

an expert for receiving the at least one parameter, the expert comprising a disk utilisation routine for calculating the disk utilisation based upon the at least one parameter.

10. A method or system according to any preceding claim, wherein the at l ast on parameter comprises:

a tim interval in seconds (T);

a number of disks in the disk array (N);

milliseconds of read time in th tim interval (MR);

milliseconds of write time in the time interval (MW);

a total number of writes in the time interval (W); and

a number of stride write destages in the time interval (S).

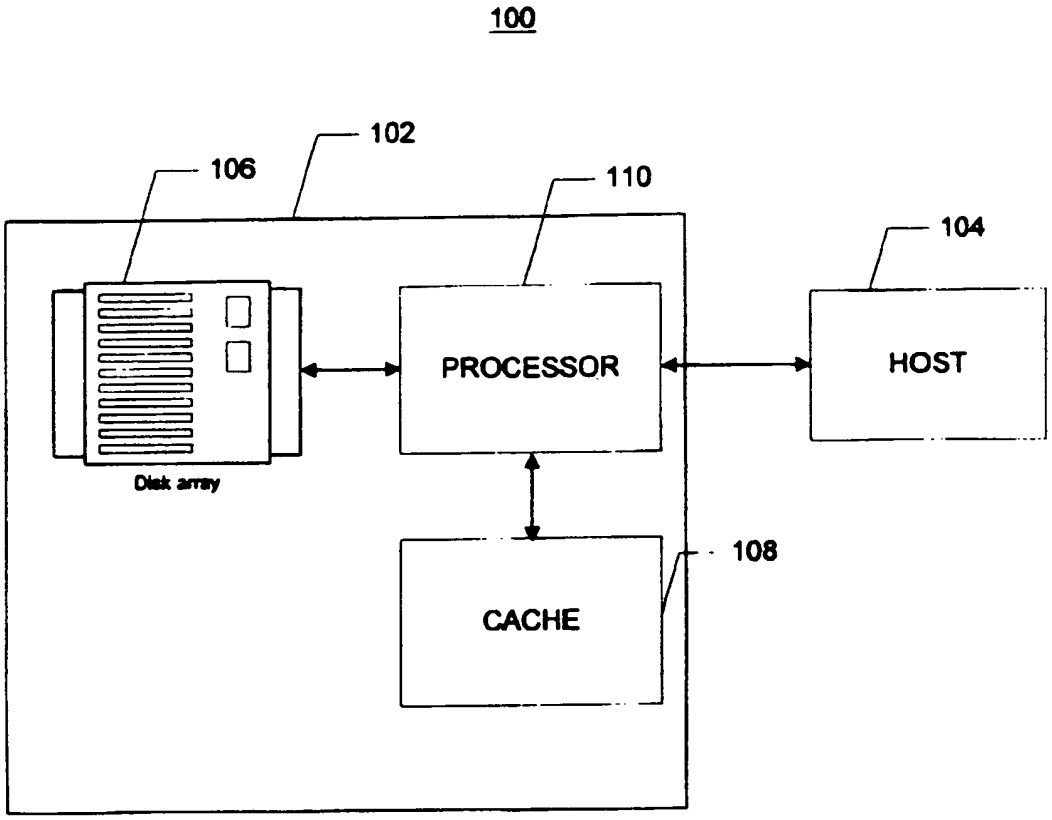
11. A method or a system according to claim 10, wherein the disk utilisation, U, is obtained from a formula U=1/(1000*T*N) * (MR + (MW/2W-S) * (4*(W-S) + N*S)).

12. A computer readable medium with computer instructions for reporting disk utilisation in a computer system, the instructions being for:

(a) collecting at least one parameter pertaining to a disk array in a disk subsystem by a processor in the disk subsystem;

(b) transferring the at least one parameter from the processor to an open system host; and

(c) calculating disk utilisation based upon the at least one parameter.



PRIOR ART

FIG. 1

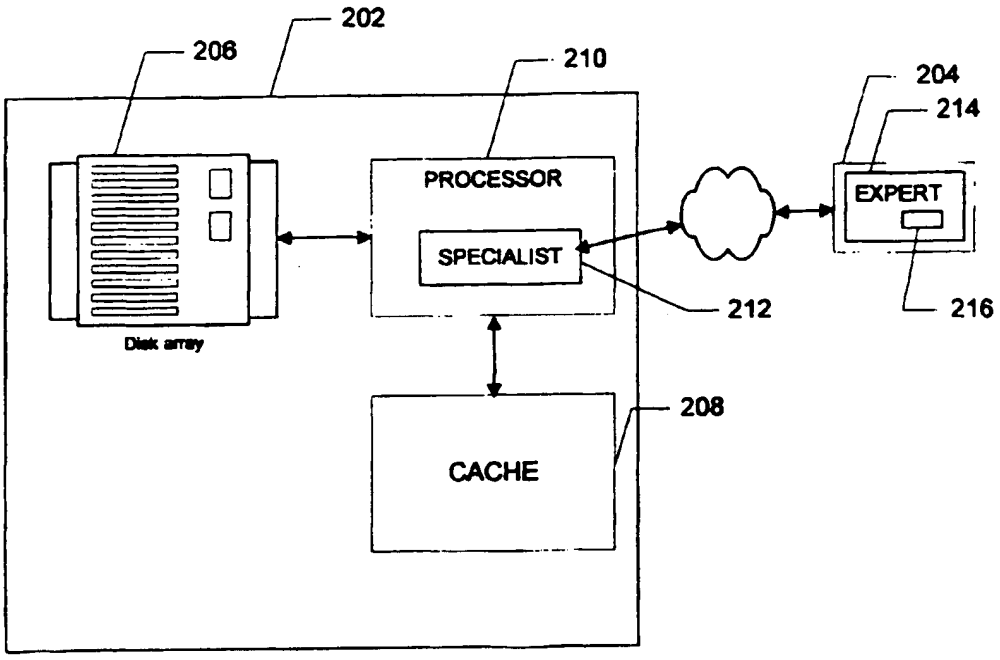

200
202
206
210
204
214
PROCESSOR
SPECIALIST
EXPERT
212
216
Disk array
208
CACHE

FIG. 2

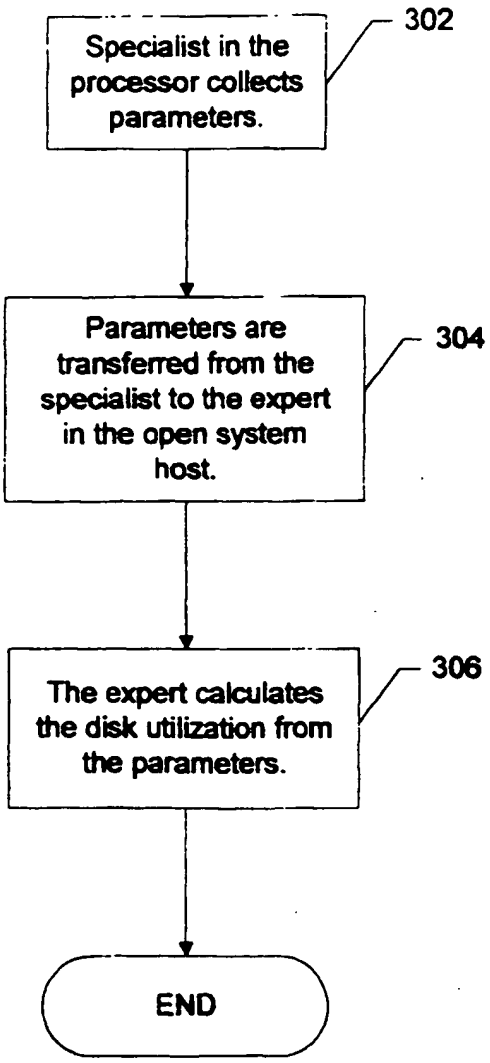

Specialist in the processor collects parameters.
302
Parameters are transferred from the specialist to the expert in the open system host.
304
The expert calculates the disk utilization from the parameters.
306
END

FIG. 3

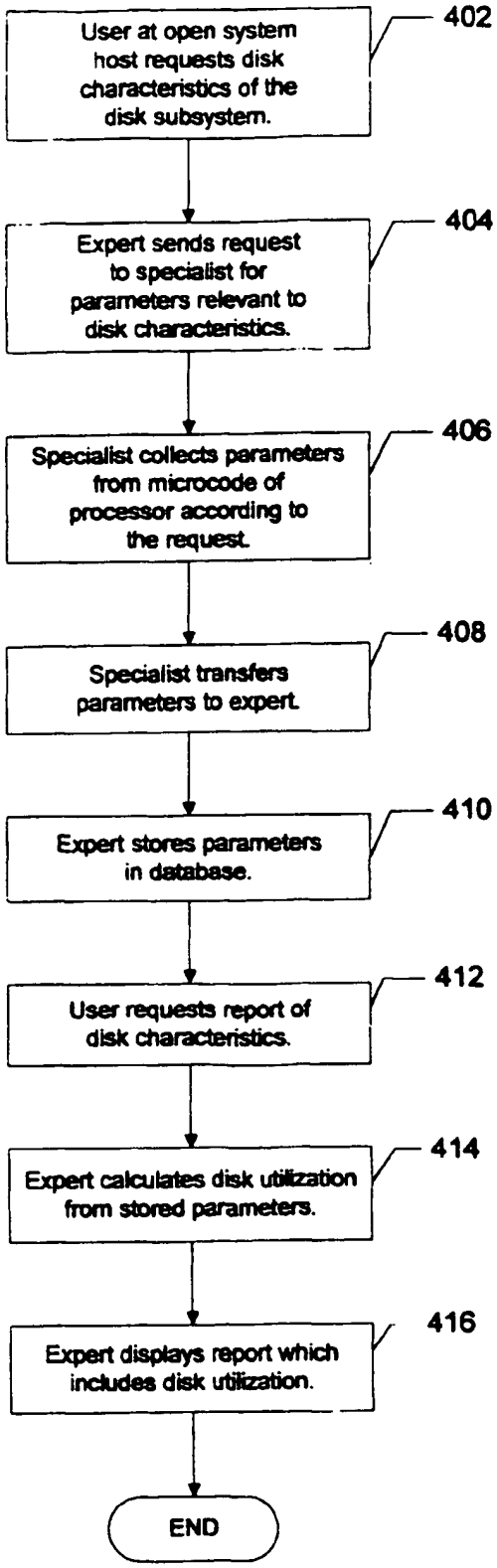

User at open system host requests disk characteristics of the disk subsystem.
402
Expert sends request to specialist for parameters relevant to disk characteristics.
404
Specialist collects parameters from microcode of processor according to the request.
406
Specialist transfers parameters to expert.
408
Expert stores parameters in database.
410
User requests report of disk characteristics.
412
Expert calculates disk utilization from stored parameters.
414
Expert displays report which includes disk utilization.
416
END

FIG. 4

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 00 30 7288

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
A	US 5 542 066 A (MATTSON ET AL.) 30 July 1996 (1996-07-30) * column 11, line 56 - line 64; figure 12 * ---	1,7,10	G06F3/06

TECHNICAL FIELDS SEARCHED (Int.Cl.7)

G06F

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
BERLIN	5 October 2000	Taylor, P

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 00 30 7288

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

05-10-2000

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5542066 A	30-07-1996	NONE	

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82